JIRSS (2006)
Vol. 5, Nos. 1-2, pp 39-52

Potential Statistical Evidence in Experiments and Renyi Information

M. Abbasnejad, N. R. Arghami

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Iran. (arghami_nr@yahoo.com)

Abstract. Recently Habibi *et al.* (2006) defined a pre-experimental criterion for the potential strength of evidence provided by an experiment, based on Kullback-Leibler distance. In this paper, we investigate the potential statistical evidence in an experiment in terms of Renyi distance and compare the potential statistical evidence in lower (upper) record values with that in the same number of iid observations from the same parent distribution.

1 Introduction and Preliminaries

Let $p(\mathbf{x})$ be the joint probability density function (pdf) of n iid observations from a distribution with pdf $f(x)$, obtained from the experiment $\mathcal{E}$, then the likelihood ratio

$$R_{\mathcal{E}}(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}$$

measures the strength of evidence in $\mathcal{E}$ favorable to the simple hypothesis $H_1 : \ p = p_1$ against the simple hypothesis $H_0 : \ p = p_0$,

Key words and phrases: Law of likelihood, record values, Renyi information, strong true evidence, weak evidence.

(Royall, 1992, 1997, 2000, 2003, Royall and Tsou, 2003).

Suppose $\mathcal{E}_1$ and $\mathcal{E}_2$ are two experiments (or sampling schemes) with (approximately) the same cost, having outcomes $\mathbf{x}$ and $\mathbf{y}$, which are the realizations of random vectors $\mathbf{X}$ and $\mathbf{Y}$, with densities $p(\mathbf{x})$ and $q(\mathbf{y})$, respectively.

When the objective of the study is to produce statistical evidence for one hypothesis against another (in the above sense), it is desirable to have a measure of performance of the experiments $\mathcal{E}_1$ and $\mathcal{E}_2$. Habibi *et al.* (2006), combining the values of potential statistical evidence, under H_1 and H_0, defined

$$S_\varphi(\mathcal{E}) = E_1\varphi[R_{\mathcal{E}}(\mathbf{X})] + E_0\varphi[1/R_{\mathcal{E}}(\mathbf{X})]$$

where $\varphi(.)$ is a non-decreasing function.
If

$$\varphi(t) = \begin{cases} 1, & t \geq K \\ 0, & t < K, \end{cases}$$

$S_\varphi(\mathcal{E})$ is the sum of the probabilities of observing strong true evidence under H_1 and H_0, where K is arbitrary and is usually between 8 and 32 (Royall, 1997).

If $\varphi(t) = t/(1+t)$, then $S_\varphi(\mathcal{E})=$ abc$(\mathcal{E})$ is the area between the cumulative distribution function (cdf) curves (under H_1 and H_0) of $\eta = R_{\mathcal{E}}(\mathbf{X})/[1 + R_{\mathcal{E}}(\mathbf{X})]$ (Emadi and Arghami, 2003).

If $\varphi(t) = \log(t)$, then

$$S_\varphi(\mathcal{E}) = D(p_1, p_0) + D(p_0, p_1) = J(p_1, p_0)$$

(Habibi *et al.*, 2006), where $D(p_1, p_0)$ and $J(p_1, p_0)$ are, respectively, asymmetric and symmetric Kullback-Leibler (K-L) distance (information) of p_1 and p_0.

The rest of the paper is organized as follows:

In Section 2 we define a pre-experimental criterion for the potential strength of evidence provided by an experiment, which is a generalization of the above criterion proposed by Habibi *et al.* (2006). In Section 3 we compare the potential statistical evidence in record values with that in the same number of iid observations from the same parent distribution, in terms of Renyi information (Renyi, 1961). In Section 4, some well known distributions are classified to three classes based on their Renyi information.

2 Renyi information as a measure of statistical evidence

For testing a general hypothesis about parameters of one population, the likelihood ratio test statistic is of general use. The likelihood ratio test statistic is a measure of deviation between the maximum likelihood achieved under the null hypothesis and the maximum achieved over the whole parameter space. Following this philosophy, a different measure of deviation, like a divergence, can be used. Some tests based on divergences have already been proposed in the literature, and it has been shown that in many cases they represent good competitors to classical tests. For example, Salicru *et al.* (1994) and Morales *et al.* (1997, 2000, 2004) suggested to test composite hypotheses, using some families of divergence, like ϕ-divergence or Renyi distance. In the following, we define Renyi distance as a pre-experimental tool.

Consider a generalization of $S_{\varphi}(\mathcal{E})$ of Section 1 to

$$S_{\varphi_1,\varphi_2}(\mathcal{E}) = \varphi_2[E_1\varphi_1[R_{\mathcal{E}}(\mathbf{X})]] + \varphi_2[E_0\varphi_1[1/R_{\mathcal{E}}(\mathbf{X})]]$$

where both $\varphi_1(.), \varphi_2(.)$ are non-decreasing or non-increasing functions.
If we take

$$\varphi_1(t) = t^{\alpha-1}$$

and

$$\varphi_2(t) = \frac{1}{\alpha-1}\log t$$

$\alpha > 0$ and $\alpha \neq 1$, then

$$\begin{aligned} S_{\varphi_1,\varphi_2}(\mathcal{E}) &= \frac{1}{\alpha-1}\log\left\{E_1\left[\frac{p_1(\mathbf{X})}{p_0(\mathbf{X})}\right]^{\alpha-1}\right\} \\ &+ \frac{1}{\alpha-1}\log\left\{E_0\left[\frac{p_0(\mathbf{X})}{p_1(\mathbf{X})}\right]^{\alpha-1}\right\} \\ &= D^{\alpha}(p_1,p_0) + D^{\alpha}(p_0,p_1) \\ &= J^{\alpha}(p_1,p_0), \end{aligned}$$

where $D^{\alpha}(p_1,p_0)$ and $J^{\alpha}(p_1,p_0)$ are, respectively, asymmetric and symmetric Renyi distance (information) of p_1 and p_0.

Example 2.1. (Bernoulli trials) Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be as follows.
$\mathcal{E}_1$: Take a random sample of size n from $B(\theta)$, a Bernoulli distribution with parameter θ.

$\mathcal{E}_2$: Continue sampling from $B(\theta)$ distribution until $p_{\theta_1}(\mathbf{x})/p_{\theta_0}(\mathbf{x}) < 1/K$ or $p_{\theta_1}(\mathbf{x})/p_{\theta_0}(\mathbf{x}) > K$.

The question may be: "Which one of $\mathcal{E}_1$ and $\mathcal{E}_2$ have more potential statistical evidence regarding the hypotheses $H_0: \ \theta = \theta_0$ and $H_1: \ \theta = \theta_1$."

Let $X = \sum_{i=1}^{n} Z_i$, where $Z_1, \ldots, Z_n$ are the outcomes of $\mathcal{E}_1$. We, then, have

$$
\begin{aligned}
S_{\varphi_1,\varphi_2}(\mathcal{E}_1) &= \frac{1}{\alpha-1}\log\left[E_{\theta_1}\left(\frac{\theta_1^{X(\alpha-1)}(1-\theta_1)^{(n-X)(\alpha-1)}}{\theta_0^{X(\alpha-1)}(1-\theta_0)^{(n-X)(\alpha-1)}}\right)\right]\\
&+ \frac{1}{\alpha-1}\log\left[E_{\theta_0}\left(\frac{\theta_0^{X(\alpha-1)}(1-\theta_0)^{(n-X)(\alpha-1)}}{\theta_1^{X(\alpha-1)}(1-\theta_1)^{(n-X)(\alpha-1)}}\right)\right]\\
&= \frac{1}{\alpha-1}\log\left[E_{\theta_1}\left((\frac{1-\theta_1}{1-\theta_0})^{n(\alpha-1)}\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)^{X(\alpha-1)}\right)\right]\\
&+ \frac{1}{\alpha-1}\log\left[E_{\theta_0}\left((\frac{1-\theta_0}{1-\theta_1})^{n(\alpha-1)}\left(\frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}\right)^{X(\alpha-1)}\right)\right]\\
&= \frac{1}{\alpha-1}\log[(\frac{1-\theta_1}{1-\theta_0})^{n(\alpha-1)}]\\
&+ \frac{1}{\alpha-1}\log\left[\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)^{\alpha-1}\theta_1+(1-\theta_1)\right]^n\\
&+ \frac{1}{\alpha-1}\log[(\frac{1-\theta_0}{1-\theta_1})^{n(\alpha-1)}]\\
&+ \frac{1}{\alpha-1}\log\left[\left(\frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}\right)^{\alpha-1}\theta_0+(1-\theta_0)\right]^n\\
&= \frac{n}{\alpha-1}\log\Big[1-\theta_0-\theta_1+2\theta_1\theta_0+\frac{(1-\theta_0)^{\alpha}\theta_1^{\alpha}}{(1-\theta_1)^{\alpha-1}\theta_0^{\alpha-1}}\\
&\quad +\frac{(1-\theta_1)^{\alpha}\theta_0^{\alpha}}{(1-\theta_0)^{\alpha-1}\theta_1^{\alpha-1}}\Big].
\end{aligned}
$$

Ignoring the "over shoot", for $\mathcal{E}_2$ we have

$$
\begin{aligned}
S_{\varphi_1,\varphi_2}(\mathcal{E}_2) &= \frac{1}{\alpha-1}\log[K^{\alpha-1}P_{\theta_1}(R>K)+(\frac{1}{K})^{\alpha-1}P_{\theta_1}(R<\frac{1}{K})]\\
&+ \frac{1}{\alpha-1}\log[(\frac{1}{K})^{\alpha-1}P_{\theta_0}(R>K)+K^{\alpha-1}P_{\theta_0}(R<\frac{1}{K})]\\
&\simeq \frac{1}{\alpha-1}\log[K^{\alpha-1}(1-\frac{1}{K+1})+(\frac{1}{K})^{\alpha-1}(\frac{1}{K+1})]
\end{aligned}
$$

$$
\begin{aligned}
&+ \quad \frac{1}{\alpha-1}\log[(\frac{1}{K})^{\alpha-1}(\frac{1}{K+1}) + K^{\alpha-1}(1-\frac{1}{K+1})] \\
&= \quad \frac{1}{\alpha-1}\log\left[\frac{K^{2\alpha-1}+1}{K^{\alpha-1}(K+1)}\right],
\end{aligned}
$$

where the approximate equality follows from Wald inequalities in the theory of SPRT (Rohatgi, 1976, pp. 616-617).

Experiments $\mathcal{E}_1$ and $\mathcal{E}_2$ will have, on average, approximately the same cost if $n = E(N)$, where N is the final sample size of $\mathcal{E}_2$.
If $\theta_0 = \frac{1}{3}$ and $\theta_1 = \frac{2}{3}$ then

$$
E_{\theta_i}(N) = \frac{[1-2/(K+1)]\log K}{(1/3)\log 2} \qquad i = 0, 1.
$$

For $K = 8$ we can see the result in Figure 1.

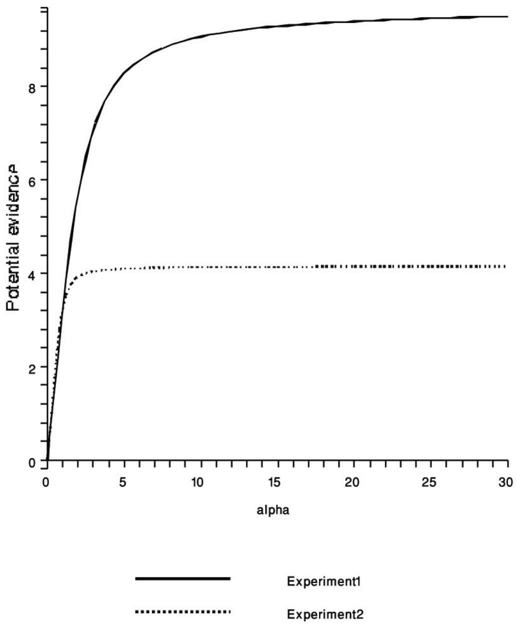

Figure 1: Potential evidence (symmetric Renyi information) of $\mathcal{E}_1$ and $\mathcal{E}_2$ in Example 2.1 for $K = 8,\ \theta_0 = 1/3$ and $\theta_1 = 2/3$.

According to Figure 1, $\mathcal{E}_1$ has more potential evidence than $\mathcal{E}_2$ for $\alpha > 1$. The reverse holds for $0 < \alpha < 1$, but the difference between $\mathcal{E}_1$ and $\mathcal{E}_2$ is not remarkable. For $0 < \alpha < 1$, $\mathcal{E}_2$ may be preferable to $\mathcal{E}_1$, because $\mathcal{E}_2$ has no probability of weak evidence (that is $P_{\theta_i}(1/K < R < K) = 0,\ i = 0, 1$). On the other hand $\mathcal{E}_1$ has the advantage of having a fixed sample size. We have similar result for $K = 16$ (See Figure 2).

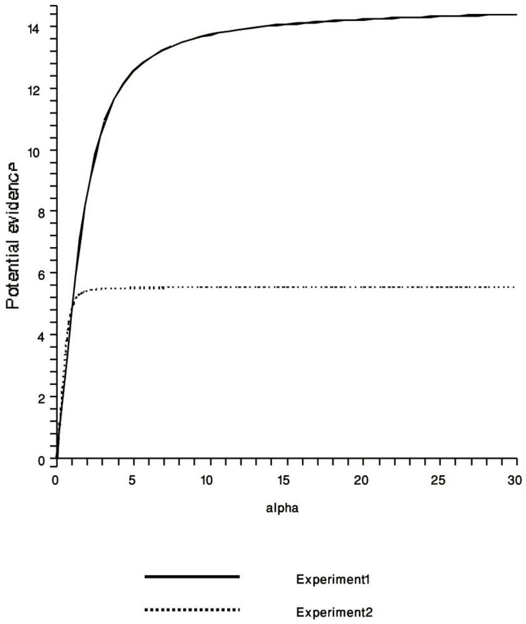

Figure 2: Potential evidence (symmetric Renyi information) of $\mathcal{E}_1$ and $\mathcal{E}_2$ in Example 2.1 for $K = 16,\ \theta_0 = 1/3$ and $\theta_1 = 2/3$.

Example 2.2. (Measuring the strength of wooden beams, Glick (1978)) Pressure is continuously increased until the beam breaks. The cost of the experiment is assumed to be equal to the number of broken beams. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be as follows.

$\mathcal{E}_1$: We measure the strength of each one of n wooden beams, so the number of broken beams is equal to n.

$\mathcal{E}_2$: We break only the beams that are weaker than all previous beams. We continue until n beams are broken, so the number of broken beams is equal to the number of lower record values.

The question is "Do the first n lower record values have more (less, equal) expected statistical true evidence as compared the same number, n, of iid observations from the same parent population?"

Ahmadi and Arghami (2001 and 2003) classified many classic families of distributions into three classes RMI, RLI and REI according to whether record values contain More, Less or Equal amount of Fisher information when compared with the same number of iid observations. Similar classification has been done by Hofmann (2004). Habibi *et al.* (2006) did a similar classification based on K-L information.

In the next Section, we first discuss the distribution of record values and then compare their potential evidence with that of *iid* observations in terms of Renyi information.

3 Record values and iid observations

Let X_i, $i \geq 1$, be a sequence of iid continuous random variables. An observation X_j will be called a lower record value if its value is smaller than that of all previous observations. Thus X_j is a lower record value if $X_j < X_i$ for all $i < j$. By convention X_1 is the first lower record value.

The times at which lower record values appear are given by the random variables T_j which are called record times and are defined by $T_1 = 1$ with probability 1 and, for $j \geq 2$, $T_j = \text{Min}\{i : \ X_i < X_{T_{j-1}}\}$. The waiting time between the i^{th} lower record value and the $(i+1)^{th}$ lower record value is called the inter-record time (IRT), and is denoted by $\Delta_i = T_{i+1} - T_i, \ i = 1, 2, \ldots$. Record times and inter-record times for upper record values are defined analogously.

Let $L_1, L_2, ..., L_n$ be the first n lower record values from a distribution with the cdf $F(x;\theta)$ and the pdf $f(x;\theta)$. Then the pdf of the joint distribution of the first n lower record values is given by

$$q(\mathbf{l};\theta) = \prod_{i=1}^{n-1} \frac{f(l_i;\theta)}{F(l_i;\theta)} f(l_n;\theta),$$

and the marginal density of L_i (the i^{th} lower record value, $i \geq 1$) is given by

$$q_i(l_i) = \frac{[-\log F(l_i;\theta)]^{i-1}}{(i-1)!} f(l_i;\theta).$$

The joint distribution of lower record values and their IRT's has density

$$q(\mathbf{l}, \boldsymbol{\Delta};\theta) = \prod_{i=1}^{n} f(l_i;\theta)[1 - F(l_i;\theta)]^{\Delta_i - 1},$$

and the joint density of L_i and Δ_i is

$$q_i(l_i, \Delta_i; \theta) = \frac{[-\log F(l_i;\theta)]^{i-1}}{(i-1)!} f(l_i;\theta)F(l_i;\theta)[1-F(l_i;\theta)]^{\Delta_i - 1}.$$

Let $U_1, U_2, ..., U_n$ be the first n upper record values from a distribution with the cdf $F(x;\theta)$ and the pdf $f(x;\theta)$. Then the pdf of the joint distribution of the first n upper record values is given by

$$q(\mathbf{u};\theta) = \prod_{i=1}^{n-1} \frac{f(u_i;\theta)}{1-F(u_i;\theta)} f(u_n;\theta),$$

and the marginal density of U_i (the i^{th} upper record value, $i \geq 1$) is given by

$$q_i(u_i) = \frac{[-\log(1-F(u_i;\theta))]^{i-1}}{(i-1)!} f(u_i;\theta).$$

The joint distribution of upper record values and their IRT's has density

$$q(\mathbf{u}, \boldsymbol{\Delta}; \theta) = \prod_{i=1}^{n} f(u_i;\theta)[F(u_i;\theta)]^{\Delta_i - 1},$$

and the joint density of U_i and Δ_i is

$$q_i(u_i, \Delta_i; \theta) = \frac{[-\log(1-F(u_i;\theta))]^{i-1}}{(i-1)!} f(u_i;\theta)(1-F(u_i;\theta))[F(l_i;\theta)]^{\Delta_i - 1}.$$

See Arnold *et al.* (1998) for more details.

In experiments such as in Example 2.2, where the experimenter has a choice between observing n iid random variables or n record values from the same distribution (almost at the same cost), it is desirable to know which experiment provides us (on average) with more statistical true evidence, that is which one of $J^{\alpha}(p_{\theta_1}, p_{\theta_0})$ or $J^{\alpha}(q_{\theta_1}, q_{\theta_0})$ is greater.

We shall call the family of distributions $\{f(x;\theta); \theta \in \Omega\}$ RMI, RLI or REI if $J^{\alpha}(q_{\theta_1}, q_{\theta_0})$ is More than, Less than or Equal to $J^{\alpha}(p_{\theta_1}, p_{\theta_0})$ respectively, where p_θ is the joint distribution of $X_1, \ldots, X_n$ and q_θ is the joint distribution of $L_1, \ldots, L_n$ or $U_1, \ldots, U_n$.

We should note that throughout this section we are considering record values without their IRT's.

Example 3.1. (Extreme value distribution, location family) The parametric family with pdf

$$f(x;\theta) = \beta e^{-\beta(x-\theta)} \exp\{-e^{-\beta(x-\theta)}\}, \ x \in \Re, \ \theta \in \Re,$$

where β is assumed to be known, is called extreme value (location) family. This family is REI, that is, lower record values contain equal amount of Renyi information when compared with the same number of iid observations from the original distribution. This is implied by Theorem 3.1 below.

We shall denote by C_1 the class of all continuous distribution functions F such that

$$F(x;\theta) = e^{-a(\theta)b(x)},$$

where $a(.)$ and $b(.)$ are real positive functions and $b(.)$ is decreasing.

This class includes, several important distributions such as:

- Extreme value distribution (location family) with cdf

$$F(x;\theta) = \exp\{-e^{-\beta(x-\theta)}\},\ x \in \Re,\ \theta \in \Re.$$

- Power distribution with cdf

$$F(x;\theta) = x^{\theta} = \exp\{-\theta \log(\frac{1}{x})\},\ 0 < x < 1;\ \theta > 0.$$

- Frechet distribution (scale family) with cdf,

$$F(x;\theta) = \exp\{-\theta x^{-\beta}\},\ x > 0;\ \theta > 0.$$

Ahmadi and Arghami (2001) showed that for all members of the class C_1 of families of distributions, L_n (the n^{th} lower record value) is a sufficient statistic for the entire set of the first n lower record values and $b(L_n)$ is distributed as $T(\mathbf{X}) = \sum_{i=1}^{n} b(X_i)$.

Theorem 3.1. *All members of the class C_1 are REI, that is, lower record values contain equal amount of Renyi information when compared with the same number of iid observations from the original distribution.*

Proof. We have

$$f(x;\theta) = a(\theta)(-b^{'}(x))F(x;\theta),$$

so

$$q(l_n;\theta) = -[a(\theta)b^{'}(l_n)/(n-1)!][a(\theta)b(l_n)]^{n-1}\exp\{-a(\theta)b(l_n)\}.$$

Thus

$$R(l_n) = [\frac{a(\theta_1)}{a(\theta_0)}]^n \exp\{-[a(\theta_1) - a(\theta_0)]b(l_n)\}.$$

Hence

$$
\begin{aligned}
D^{\alpha}(q_{\theta_1}, q_{\theta_0}) &= \frac{1}{\alpha-1}\log\Big\{E_{\theta_1}[(R(L_n))^{\alpha-1}]\Big\} \\
&= \frac{1}{\alpha-1}\log\Big\{[\frac{a(\theta_1)}{a(\theta_0)}]^{n(\alpha-1)} \\
&\quad \times E_{\theta_1}\Big[\exp[-(\alpha-1)(a(\theta_1)-a(\theta_0))b(l_n)]\Big]\Big\} \\
&= n\log[\frac{a(\theta_1)}{a(\theta_0)}] \\
&+ \frac{1}{\alpha-1}\log E_{\theta_1}\Big[\exp(-(\alpha-1)(a(\theta_1)-a(\theta_0))\sum_{i=1}^{n} b(X_i)\Big] \\
&= D^{\alpha}(p_{\theta_1}, p_{\theta_0}). \qquad \square
\end{aligned}
$$

Example 3.2. (Weibull distribution) The parametric family with pdf

$$f(x;\theta) = \theta\beta x^{\beta-1}e^{-\theta x^{\beta}},\ x>0,\ \theta,\beta>0,$$

is called weibull (scale) family, β is assumed to be known. This family is REI, that is, upper record values contain equal amount of Renyi information when compared with the same number of iid observations from the original distribution. This is implied by Theorem 3.2 below.

We shall denote by C_2 the class of all continuous distribution functions F such that

$$F(x;\theta) = 1 - e^{-a(\theta)b(x)},$$

where $a(.)$ and $b(.)$ are real positive functions and $b(.)$ is increasing.

This class includes, several important distributions such as:
- Weibull distribution (location family) with cdf

$$F(x;\theta) = 1 - e^{-\theta x^{\beta}},\ x>0,\ \theta,\beta>0.$$

- Parto distribution with cdf

$$F(x;\theta) = 1 - \frac{1}{x^{\theta}} = 1 - e^{-\theta\ln x},\ x>1;\ \theta>0.$$

- Exponential distribution with cdf

$$F(x;\theta) = 1 - e^{-\theta x},\ x>0;\ \theta>0.$$

Ahmadi and Arghami (2001) showed that for all members of the class C_2 of families of distributions, U_n (the n^{th} upper record value) is a sufficient statistic for the entire set of the record values and $b(U_n)$ is distributed as $T(\mathbf{X}) = \sum_{i=1}^{n} b(X_i)$.

Theorem 3.2. *All members of the class C_2 are REI, that is, upper record values contain equal amount of Renyi information when compared with the same number of iid observations from the original distribution.*

The proof is similar to that of Theorem 3.1 and is thus omitted.

In Section 4 we show that a theorem, proved in Abbasnejad and Arghami (2006), can be used to extend the classifications of families by Ahmadi and Arghami (2001 and 2003) and Habibi *et al.* (2006), to the case of Renyi information.

4 Classification of some well known distribution

In the previous section we have discussed two classes of family of distributions characterized by the sufficiency of upper and lower record values. Ahmadi and Arghami (2001 and 2003) classified many classic families of distributions into REI, RMI and RLI classes based on Fisher information. Habibi *et al.* (2006) did the same thing based on Kullback-Leibler distance. Our aim here is to extend their results based on Renyi information by using the following theorem. We derive Table 4.1 below from Table II of Ahmadi and Arghami (2003).

Theorem 4.1. (Theorem 4.1 of Abbasnejad and Arghami (2006)) *Let $\{p_\theta,\ \theta \in \Omega\}$ and $\{q_\theta,\ \theta \in \Omega\}$ be two families of densities. Assume that both family have finite Fisher and Renyi information, continuous in their arguments. If the following conditions hold*

(i) $I_X(\theta) - I_Y(\theta) \geq d_0 > 0,\ \forall\theta \in I = [\theta_0, \theta_1]$

(ii) The third derivatives of $D_X^\alpha(p_{\theta+\delta}, p_\theta)$ and $D_Y^\alpha(q_{\theta+\delta}, q_\theta)$ with respect to (w.r.t) δ are bounded for every $\theta \in I$ and every δ in the neighborhood $I_0 = [0, c]$ of zero, then

$$D_X^\alpha(p_{\theta_1}, p_{\theta_0}) > D_Y^\alpha(q_{\theta_1}, q_{\theta_0}).$$

Table 4.1 below, reproduced from Ahmadi and Arghami (2003), indicates the types (RMI, REI or RLI) of a number of important families of distributions with respect to Fisher (and thus, by Theorem 4.1, wrt Renyi) information.

Table 4.1. Families of distributions classified on the basis of Renyi information. (reproduced from Ahmadi and Arghami (2003), RMI cases are printed in bold face.)

cdf	Upper record without IRT	Upper record with IRT	Lower record without IRT	Lower record with IRT
$N(\theta, \sigma^2)$	RLI	**RMI**	RLI	**RMI**
$N(\beta, \theta)$	**RMI**	**RMI**	**RMI**	**RMI**
$\Gamma(\beta, \theta),\ \ 0 < \beta < 1$	**RMI**	**RMI**	RLI	RLI
$\Gamma(\beta, \theta),\ \ \beta = 1$	REI	**RMI**	RLI	REI
$\Gamma(\beta, \theta),\ \ \beta > 1$	RLI	**RMI**	RLI	**RMI**
$1 - \exp(-\theta x)$	REI	**RMI**	RLI	REI
x^θ	RLI	REI	REI	**RMI**
$\exp\{-\exp[-\beta(x-\theta)]\}$	RLI	REI	REI	**RMI**
$\exp\{-\exp[-\theta(x-\beta)]\}$	REI	RLI	**RMI**	REI
$L(\theta, \beta)$	RLI	**RMI**	RLI	**RMI**
$L(\beta, \theta)$	RLI	**RMI**	RLI	**RMI**
$1 - x^{-\theta}$	REI	**RMI**	RLI	REI
$1 - \exp(-\frac{1}{2\theta^2}x^2)$	REI	**RMI**	RLI	REI
$1 - (1 + x^\theta)^{-\beta}$	REI	**RMI**	RLI	REI
$1 - \exp(-\theta x^\beta)$	REI	**RMI**	RLI	REI

Our classification is also consistent with that of Habibi *et al.* (2006).

References

Abbasnejad, M. and Arghami, N. R. (2006), Relations between Renyi and Fisher information. Submitted.

Ahmadi, J. and Arghami, N. R. (2001), On the Fisher information in record values. Metrika, **53**, 195-205.

Ahmadi, J. and Arghami, N. R. (2003), Comparing the Fisher information in record values and IID observations. Statistics, **37**, 435-441.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998), Records. New York: John Wiley.

Emadi, M. and Arghami, N. R. (2003), Some measures of support for statistical hypotheses. J. Stat. Theory Appl., **2**, 165-176.

Glick, N. (1978), Breaking record and breaking boards. Amer. Math. Monthly, **85**, 2-26.

Habibi, A., Ahmadi, J. and Arghmi, N. R. (2006), Statistical evidence in experiments and in record values. Communication in Statistics, Theory and Methods., **35**, 1971-1983.

Hofmann, G. (2004), Comparing Fisher information in record data and random observations. Statistical Papers, **45**, 517-528.

Kullback, S. (1959), Information Theory and Statistics. New York: John Wiley.

Morales, D., Pardo, L., and Vajda, I. (1997), Some new statistics for testing hypotheses in parametric models. Multivariate Anal., **62**, 137-168.

Morales, D., Pardo, L., and Vajda, I. (2000), Renyi statistic in directed families of exponential experiments. Statistics, **34**, 151-174.

Morales, D., Pardo, L., Pardo, M. C., and Vajda, I. (2004), Renyi statistics for testing composite hypotheses in general exponential models. Statistics, **38**, 133-147.

Renyi, A. (1961), On measures of entropy and information. Proc. Fourth. Berkeley Symp. Math. Stat. Prob. 1960, Vol I, Berkeley: University of California Press, 547-561.

Rohatgi, V. K. (1976), An Introduction to Probability Theory and Mathematical Statistics. New York: John Wiley.

Royall, R. M. (1992), The elusive concept of statistical evidence. Bayesian Statistics 4, Eds. Bernardo et al., 405-413.

Royall, R. M. (1997), Statistical Evidence: A Likelihood Paradigm. Chapman and Hall.

Royall, R. M. (2000), On the probability of observing misleading statistical evidence. J. Amer. Statist. Assoc., **95**, 760-780.

Royall, R. M. (2003), Interpreting statistical evidence about a finite population. Analysis of survey data., Wiley Ser. sur. Methodal., Chichester: Wiley, 59-71.

Royall, R. M. and Tsou, T. (2003), Interpreting statistical evidence by using imperfect models: Robust adjusted likelihood functions. J. R. Stat. Soc., **65**, 391-404.

Salicru, M., Morales, D., Mendenenz, M. L., and Pardo, L. (1994), On the applications of divergence type measures in testing statistical hypotheses. Multivariate Anal., **51**, 372-391.